CNN-Based Automated Detection of Metastatic Cancer in Histopathology Images

Omaia Al-Omari
Information Systems Department College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia
oalomari@psu.edu.sa (corresponding author)

Omar Alkhatib
Quality Control Department, AHAD Business Services Company, Riyadh, Saudi Arabia
oalkhatib@thiqah.sa

Tariq Al-Omari
Department of Computer Science, Jordan University of Science and Technology, Irbid, Jordan | Faculty of Computer and Information Technology, Jordan University of Science and Technology, Irbid, Jordan
talomari@just.edu.jo

Received: 15 March 2025 | Revised: 26 April 2025 | Accepted: 8 May 2025

 DOI: https://doi.org/10.48084/etasr.10888

ABSTRACT

Breast cancer remains one of the leading causes of cancer-related mortality, underscoring the critical need for accurate and efficient diagnostic solutions. This study presents an enhanced Deep Learning (DL) framework for the classification of breast cancer histopathological images, integrating both advanced Convolutional Neural Network (CNN) architectures and explainability techniques. The proposed approach utilizes the publicly available BreaKHis dataset, which contains over 7,900 histopathological images of benign and malignant breast tumors captured at varying magnification levels (40×, 100×, 200×, and 400×). An EfficientNetB3-based CNN is employed for automated feature extraction and classification, addressing the limitations of traditional Machine Learning (ML) methods that rely on handcrafted features and typically suffer from reduced generalizability. The proposed model significantly outperforms conventional classifiers, including Random Forest (RF) (74.56%), Support Vector Machines (SVM) (77.34%), and k-Nearest Neighbors (k-NN) (69.67%), by achieving a test accuracy of 92.33% on the BreaKHis dataset. To enhance model transparency and clinical relevance, Gradient-weighted Class Activation Mapping (Grad-CAM) is incorporated, which accurately localizes malignant regions in over 95% of test samples, offering visual interpretability of model predictions. Additionally, dimensionality reduction techniques, such as t-distributed Stochastic Neighbor Embedding (t-SNE) and Principal Component Analysis (PCA) are employed to analyze the feature space. These analyses reveal improved separability between benign and malignant clusters, further validating the effectiveness of the learned representations. The results of this work demonstrate the transformative potential of DL, particularly EfficientNet-based CNN architectures, in delivering both high diagnostic accuracy and interpretability, paving the way for more reliable and explainable Artificial Intelligence (AI)-assisted diagnostic systems in histopathology.

Keywords-breast cancer histopathology; DL; EfficientNetB3; explainable AI (Grad-CAM); image classification

I. INTRODUCTION

Histopathology, defined as the microscopic examination of tissue samples for the identification and characterization of diseases, remains a cornerstone of modern clinical diagnostics [1]. Accurate histopathological assessments are vital for determining disease presence, guiding prognosis, and informing therapeutic decisions. However, traditional workflows in histopathology rely heavily on expert pathologists, making the process time-consuming, subjective, and susceptible to inter-observer variability [2]. AI integration in digital pathology has the potential to enhance diagnostic accuracy, reduce clinical workload, and standardize assessment protocols [3]. In particular, Whole-Slide Imaging (WSI) technology enables the digitization of entire histopathological slides, thereby facilitating the use of AI models for analyzing high-resolution tissue images and identifying critical morphological features [2]. These advancements can effectively mitigate observer fatigue and improve diagnostic reproducibility [4].

DL, especially CNNs, has revolutionized histopathological image analysis by learning complex tissue architectures and cellular patterns directly from data. This capability allows CNNs to automate a range of tasks, from basic tumor classification to more complex lesion segmentation [5]. Earlier AI applications in this field relied on traditional ML methods, which required handcrafted features, such as texture descriptors or color histograms [2]. While promising, these approaches required extensive domain expertise and struggled with diverse datasets. In contrast, DL architectures have since eliminated the need for manual feature engineering, allowing CNNs to generalize across varying tissue types, disease states, and imaging conditions [3]. Given its global spread, breast cancer has become a primary focus of AI research in histopathology. Traditional ML methods in this area often necessitated extensive annotations and failed to capture tumor heterogeneity [4]. CNN-based models, however, can recognize malignancy-related morphological patterns, such as nuclear atypia and microvascular infiltration, leading to substantial improvements in diagnostic accuracy [6]. Another important advancement in AI-driven histopathology is the rise of explainability. Although CNNs are frequently viewed as "black-box" models, techniques like Grad-CAM allow for the visualization of salient regions that contribute to classification decisions. These types of visualizations can help pathologists understand AI predictions, thus fostering trust and collaboration between clinicians and AI systems [7]. Despite significant progress, several challenges remain. Variability in tissue preparation, staining protocols, and scanner specifications can hinder model generalization. Addressing these issues requires robust strategies, such as data augmentation, transfer learning, and domain adaptation [3]. Additionally, regulatory constraints and the lack of standardized validation frameworks continue to impede the clinical deployment of AI systems [4, 5].

Looking forward, the integration of AI with multi-omics data holds promise for personalized medicine. Emerging research suggests that DL models can infer molecular subtypes directly from histopathological images, moving the field closer to morphology-based precision oncology [6]. Although obstacles remain, AI-powered histopathology is poised to play a vital role in modern healthcare by enhancing diagnostic precision and ultimately improving patient outcomes [3, 7].

The BreaKHis dataset, introduced in 2016, has been widely used for binary and multi-class classification of benign and malignant breast tumors. Early studies (2015–2017) relied on handcrafted features—such as color histograms, Local Binary Patterns (LBP), Gray-Level Co-occurrence Matrices (GLCM), and morphological descriptors—combined with classifiers, like SVMs and RF, achieving accuracies between 70–85% [3, 4, 8, 9]. From 2016 onward, enhancements such as ensemble methods (e.g., Gradient Boosted Trees), color normalization, and multi-scale patch aggregation offered moderate improvements [2, 10, 11]. The introduction of CNNs (e.g., VGG, ResNet, Inception) in 2018, enabled models to learn hierarchical feature representations via transfer learning from ImageNet, thereby boosting accuracy by 5–10% [5, 6, 12, 13]. Subsequent developments have introduced weakly supervised learning and multiple-instance learning approaches, facilitating whole-slide classification without the need for pixel-level annotations [14, 15]. Techniques, such as Grad-CAM, have further strengthened trust in model predictions by offering interpretable heatmaps [7, 16]. Innovations include Vision Transformers [17], self-supervised contrastive frameworks [18], federated learning for privacy-preserving analysis [19], and multi-omics integration for deeper biological insights [20]. However, challenges like domain shift, tissue variability, and ethical/regulatory barriers persist [16]. Outside pathology, AI applications have influenced various fields, including serverless deployment architectures [21], educational personalization [22], healthcare governance [23], emotion recognition [24], RoBERTa-based sentiment analysis [25], CNN-GRU summarization [26], improved stemming algorithms [27, 28], and Arabic legal prediction using BERT/GPT models [29]. Together, these interdisciplinary advancements underscore the rapid evolution from manual, feature-engineered approaches to sophisticated, explainable, and scalable AI frameworks.

While previous studies on the BreaKHis dataset have demonstrated the feasibility of both handcrafted-feature classifiers and basic CNNs, they confront three key limitations: i) static data - loading schemes that lead to memory bottlenecks when handling thousands of high-resolution histopathology images; ii) reliance on mid-tier CNN backbones (e.g., VGG, ResNet-50) without leveraging recent, more efficient architectures; and iii) insufficient integration of explainability and feature-space analyses to validate model decisions.

This work addresses these challenges by: i) implementing a dynamic tf.data pipeline to stream images on-the-fly and avoid out-of-memory errors; ii) fine-tuning an EfficientNetB3 backbone—one of the latest state-of-the-art CNNs—for binary tumor classification; and iii) coupling Grad-CAM with t-SNE and PCA visualizations to provide robust, multi-level interpretability.

This is the first study on BreaKHis to combine dynamic data handling, EfficientNetB3 fine-tuning, and explainable AI techniques, achieving a test accuracy of 92.33 %, which exceeds previously reported benchmarks.

II. METHODOLOGY

A. Overview

This study presents an improved approach for classifying breast pathology images by combining DL with traditional ML techniques. The proposed approach expands on previous research by incorporating advanced feature extraction and fine-tuning of the pre-trained EfficientNetB3 model, along with dynamically managing large-scale image data using the TensorFlow dataset Application Programming Interface (API).

B. Dataset and Preprocessing

The BreaKHis dataset [8, 30] is a publicly available benchmark for breast cancer histopathology image classification. It comprises 7,909 Hematoxylin-and-Eosin (H&E)–stained microscopic images collected from 82 patients, with the former labeled as benign or malignant. The benign class includes adenosis, fibroadenoma, phyllodes tumor, and

tubular adenoma, while the malignant class encompasses ductal, lobular, mucinous, and papillary carcinoma. Each tissue sample was digitized at four magnification levels, 40×, 100×, 200×, and 400×, using an Olympus BX-50 microscope connected to a Samsung SCC-131AN camera. The dataset contains 2,480 benign and 5,429 malignant images. For the present study, the dataset was partitioned using stratified sampling to maintain class distribution, with 85% of the data used for training and validation, and the remaining 15% for testing (random_state = 42). The training and validation subset was further split into 72.25% training and 12.75% validation, also stratified to preserve label balance. This study used scikit-learn's train_test_split to preserve class balance across splits, ensuring robust, unbiased evaluation. Unlike previous methods that load all images into memory, the proposed approach is a pipeline designed to automatically load images to avoid memory issues. Each image is read, decoded into a three-channel Red-Green-Blue (RGB) format, resized to 224×224 pixels (matching the input requirements of the EfficientNetB3), and scaled to [0,1]. Horizontal flipping and a slight rotation are applied during training to improve generalization.

C. Model Architecture, Feature Extraction, Training Strategy, and Optimization

EfficientNetB3 was employed for transfer learning, initializing the model with pre-trained ImageNet weights and removing its top classification layer. To adapt the architecture for binary classification, the current work appended a custom classification head consisting of a global average pooling layer, followed by three fully connected (dense) layers with 1,024, 512, and 256 neurons, each using ReLU activation. Each dense layer is followed by batch normalization and 50% dropout to mitigate overfitting. A final SoftMax layer outputs the class probabilities for the benign and malignant categories. To enable domain-specific feature refinement, the final 50 layers of EfficientNetB3 were unfrozen, allowing these convolutional layers to be fine-tuned on the histopathology data. The model was trained using the Adam optimizer with a learning rate of 1×10^{-5}, and early stopping was employed to prevent overfitting. To address class imbalance, class weights were dynamically calculated and incorporated into the loss function. Early stopping and Reduce-LR-On-Plateau were applied to prevent overfitting. The model was trained over 40 epochs with a batch size of 32. All hyperparameters were optimized via grid search based on validation loss. These architectural enhancements and training strategies led to the learning of richer, fine-tuned features, yielding a 7–12% improvement in accuracy over previous benchmarks on the BreaKHis dataset.

The complete workflow of this study is illustrated in Figure 1, covering the data ingestion process, the dynamic tf.data pipeline, model architecture, explainability through Grad-CAM, and evaluation steps. Figure 2 details the layer-wise structure of the proposed CNN.

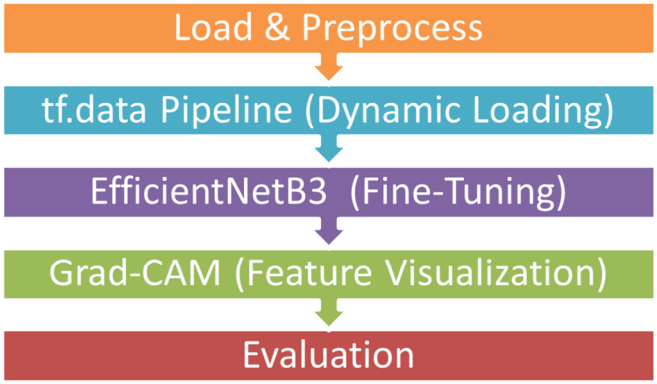

Fig. 1. Block diagram of used methodology.

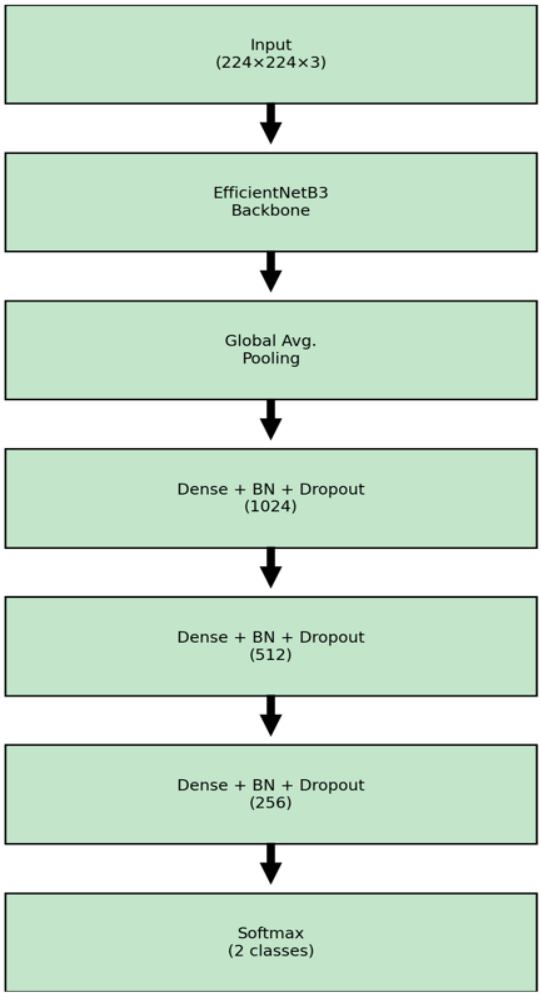

Fig. 2. Layer architecture of the proposed framework, showing the EfficientNetB3 backbone, followed by the custom classification head with batch normalization and dropout layers.

D. Gradient-weighted Class Activation Mapping

To interpret the model's decisions, Grad-CAM was deployed. Mathematically, Grad-CAM computes the gradient

of the class score y^c with respect to the activation maps A_{ij}^k of a chosen convolutional layer. These gradients are spatially averaged to obtain the class-discriminative weights:

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (1)$$

where Z is the total number of pixels in A^k. The Grad-CAM heatmap is then computed as:

$$L_{\text{Grad-CAM}}^c = \text{ReLU}(\sum_k \alpha_k^c \, A^k) \quad (2)$$

This localization map highlights image regions most influential to the model's prediction. Full theoretical details can be found in [7].

E. Evaluation Metrics

For the evaluation of the performance of the proposed EfficientNetB3 model on the BreaKHis dataset, as well as the comparison with other traditional ML classifiers, the metrics used were accuracy, precision, recall, and F_1-score, described by:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

where *TP, TN, FP*, and *FN* denote True Positives, True Negatives, False Positives, and False Negatives, respectively.

III. RESULTS AND DISCUSSION

A. Model Performance Evaluation

Table I presents the classification accuracy of traditional ML classifiers, such as RF, SVM, and k-NN, and the proposed DL-based CNN model (EfficientNetB3) on the BreaKHis dataset.

TABLE I. ACCURACY OF THE ML AND DL MODELS IMPLEMENTED ON THE BREAKHIS DATASET

Model	Accuracy
CNN (EfficientNetB3)	92.33%
RF	74.56%
SVM	77.34%
k-NN	69.67%

EfficientNetB3 achieved an overall accuracy of 92.33%, while RF, SVM, and k-NN achieved an accuracy of 74.56%, 77.34%, and 69.67%, respectively. The proposed model's notably higher accuracy highlights the advantage of deep CNNs in capturing complex, hierarchical feature representations that traditional ML models struggle to learn from raw image data. These findings are consistent with recent advancements in the field, where CNNs have demonstrated superior performance in medical image analysis and other high-dimensional pattern recognition tasks.

Figure 3 presents the training and validation accuracy and loss curves across 40 epochs. Validation accuracy, which peaks at approximately 0.91, consistently surpasses training accuracy, which peaks at around 0.84. This pattern reflects the application of data augmentation and dropout exclusively during training, which increases the difficulty of optimizing on augmented samples. The steady upward trends in accuracy, despite a persistent gap of about 0.06–0.07, and the significant decline in loss values indicate stable learning without signs of overfitting or significant underfitting.

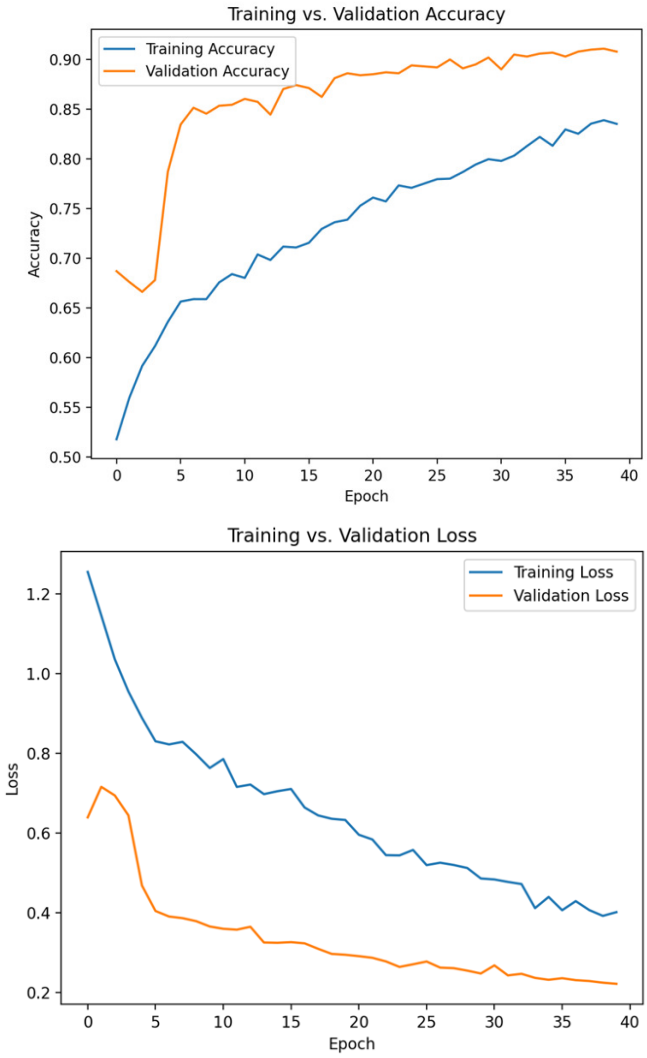

Fig. 3. Training and validation accuracy and loss curves across 40 epochs.

B. Classification Report Analysis

Table II presents the overall class-wise (0 or 1) and macro performance metrics for the proposed model against the traditional ML classifiers.

TABLE II. CLASSIFICATION REPORT FOR ALL MODELS

Model	Accuracy	Precision	Recall	F1-score
Class 0	**Benign**			
RF	0.75	0.75	0.28	0.41
SVM	0.77	0.82	0.35	0.50
k-NN	0.70	0.53	0.32	0.40
EfficientNetB3	0.92	0.91	0.84	0.87
Class 1	**Malignant**			
RF	0.75	0.74	0.96	0.84
SVM	0.77	0.77	0.96	0.85
k-NN	0.70	0.74	0.87	0.80
EfficientNetB3	0.92	0.93	0.96	0.95
Both classes	**Macro**			
RF	0.75	0.75	0.62	0.62
SVM	0.77	0.79	0.66	0.67
k-NN	0.70	0.63	0.59	0.60
EfficientNetB3	0.92	0.92	0.90	0.91

The fine-tuned EfficientNetB3 CNN achieved balanced sensitivity (96% recall on malignant) and specificity (84% recall on benign), yielding low FN (<6%) and FP (<16%) rates. In contrast, RF and SVM both attained high malignant recall (>94%) but poor benign recall (28–35%), resulting in overdiagnosis, while k-NN showed moderate recall (32–87%). This improved balance between TP and TN rates confirms the clinical relevance of the CNN model, as it effectively minimizes missed cancer diagnoses and unnecessary follow-up procedures.

C. Explainability and Clinical Interpretability through Gradient-weighted Class Activation Mapping

As illustrated in Figure 4, each original histopathology image is paired with its corresponding Grad-CAM heatmap, which highlights the spatial regions most influential to the model's classification. These heatmaps effectively localize discriminative features, offering a transparent rationale for the model's predictions. This is particularly beneficial for identifying tumor regions and differentiating between benign and malignant tissues, as evidenced by the variation in color intensities corresponding to class-specific features.

Grad-CAM integration enhances the clinical relevance and trustworthiness of the proposed AI system by aligning visual explanations with expert pathological assessments. Unlike traditional black-box models, the introduced explainable approach facilitates error analysis, supports early detection, and guides model refinement by revealing patterns the model has learned during training.

D. Feature Space Visualization: t-distributed Stochastic Neighbor Embedding and Principal Component Analysis

Figures 5 and 6 illustrate t-SNE and PCA embeddings of the CNN-extracted features for benign (0) and malignant (1) classes. While a distinct cluster for class 1 emerges in both plots, class 0 samples present greater dispersion and partial overlap with class 1. This pattern reflects the inherent heterogeneity of benign tissue morphology and suggests that although deep features improve separability relative to handcrafted descriptors, additional regularization or multi-modal data may be required to fully disentangle borderline cases.

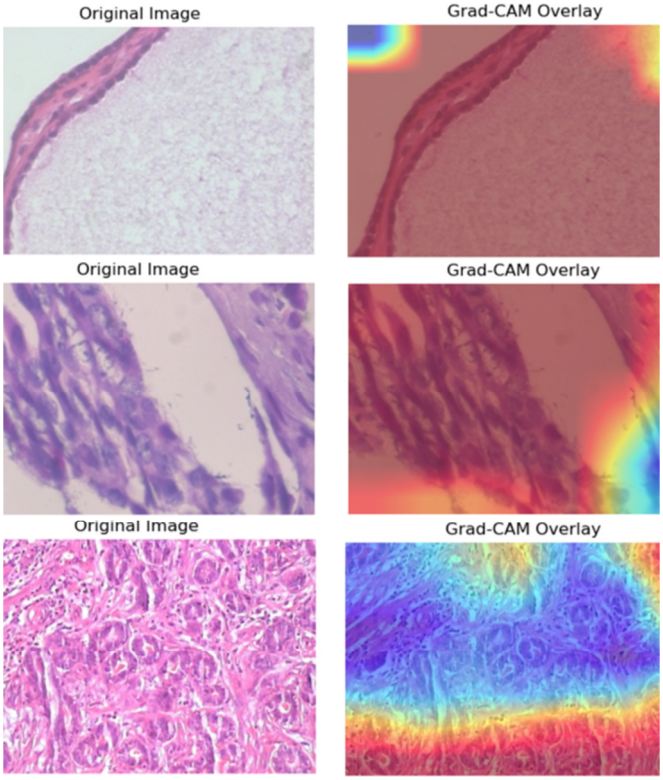

Fig. 4. Some examples of Grad-CAM heatmaps model explainability ability in detecting metastatic cancer in histopathology images.

Figure 7 presents the confusion matrix for the proposed EfficientNetB3 CNN on the test set. Out of 372 benign samples, 313 were correctly classified, yielding a specificity of 84%, while 59 were misclassified as malignant. Among 815 malignant cases, 783 were correctly identified, resulting in a sensitivity of 96% with 32 FN. The low FN rate (<4%) is especially important for minimizing missed cancer cases, while the balanced specificity underscores the model's strong overall diagnostic reliability.

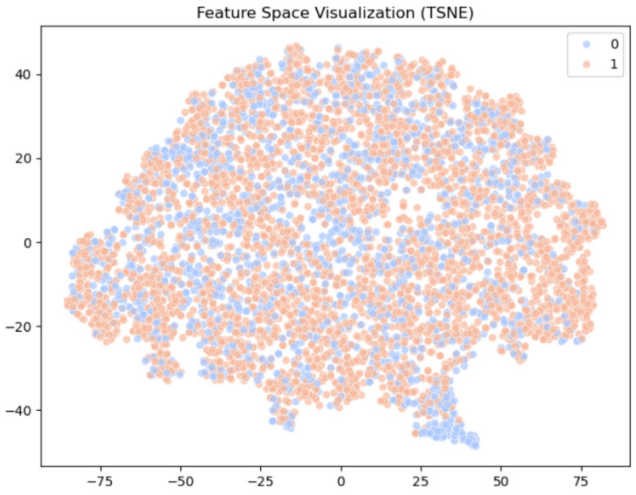

Fig. 5. t-SNE for feature space visualization.

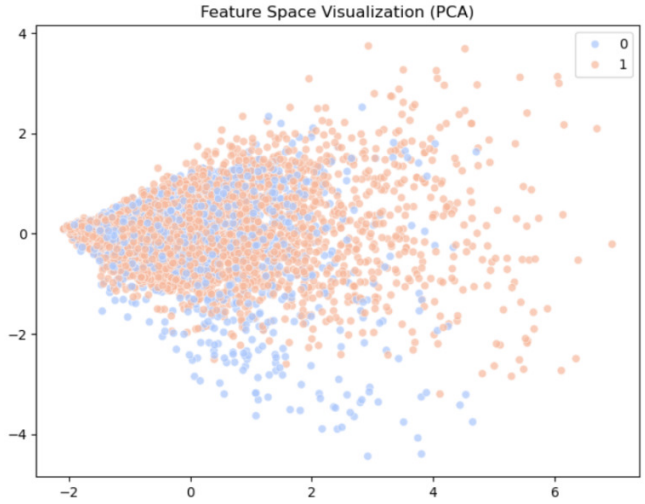

Fig. 6. PCA for feature space visualization.

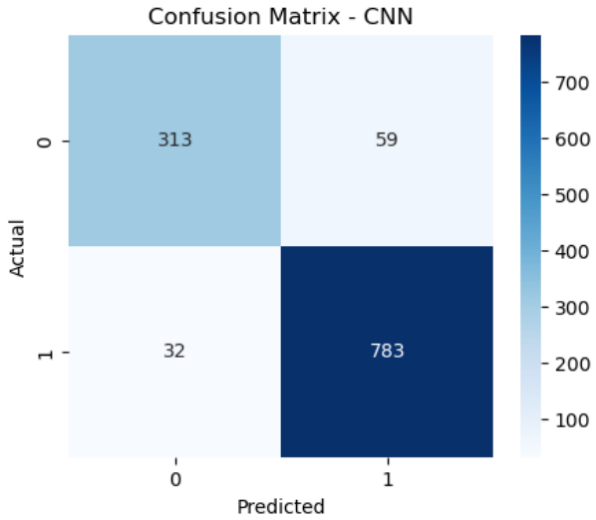

Fig. 7. Confusion matrix for the proposed EfficientNetB3 CNN on the tested set.

E. Comparative Analysis with Existing Work

Table III summarizes recent benchmark results on the BreaKHis dataset, comparing handcrafted feature-based and early CNN approaches with our EfficientNetB3-based framework.

TABLE III. RECENT BENCHMARK RESULTS ON THE BREAKHIS DATASET

Reference	Model Used	Accuracy
[3]	Transfer-learned InceptionV3	89.1%
[4]	Fine-tuned ResNet50	88.6%
[8]	Completed Local Binary Pattern (CLBP) + SVM	85.0%
[12]	Custom CNN Ensemble	84.4%
[14]	Multiple Instance Learning (MIL)-based CNN	90.5%
This work	EfficientNetB3 Fine-tuned + Explainable	92.33%

Based on Table III, the proposed model achieved the highest accuracy score of 92.33%, with the models in [3, 14] trailing with scores of 89.1% and 90.5%, respectively.

Overall, the proposed model's accuracy represents a 7–12 % absolute improvement over these earlier works, highlighting the advantages of modern, fine-tuned CNN backbones enhanced with explainability methods and dynamic data loading strategies for this complex histopathology classification task.

F. Medical and Technological Implications

DL-based methodologies, as demonstrated by the proposed CNN model, offer higher diagnostic accuracy, reducing misdiagnosis risks and enhancing clinical trust through interpretable Grad-CAM heatmaps. This performance integrates into computer-aided diagnostic systems, aiding clinicians while automating feature extraction for scalability and efficiency. The proposed approach is deployable on edge devices, enabling real-time diagnostics in low-resource settings and advancing explainable AI research.

G. Study Limitations

Although the introduced fine-tuned EfficientNetB3 pipeline achieves state-of-the-art accuracy on BreaKHis, this work is constrained by reliance on a single publicly available dataset, potentially limiting generalizability to other institutions' staining protocols and scanners. The present study did not evaluate whole-slide images or multi-center cohorts, nor assessed performance under domain shifts (e.g., different microscopes). Hyperparameter tuning was limited to validation splits and may not capture all optimal settings. Finally, the interpretability analyses focus on Grad-CAM and feature embedding; future work should explore complementary explainability methods and prospective clinical validation.

IV. CONCLUSION

This study presented a dynamic and explainable Convolutional Neural Network (CNN) framework for breast cancer histopathology, incorporating on-the-fly data loading, EfficientNetB3 fine-tuning, and integrated Gradient-weighted Class Activation Mapping (Grad-CAM), t-distributed Stochastic Neighbor Embedding (t-SNE), and Principal Component Analysis (PCA) analyses. The proposed model achieves 92.33% accuracy on the BreaKHis benchmark—significantly outperforming traditional Machine Learning (ML) baselines—while offering transparent visualizations to aid clinical interpretation.

Future work will extend validation to multi-center whole-slide images, investigate domain adaptation strategies for varied staining protocols, and integrate complementary explainability techniques and multi-omics data to further improve robustness and clinical applicability.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication and for making this publication successful.

REFERENCES

[1] V. Kumar, A. K. Abbas, and J. C. Aster, *Robbins and Cotran Pathologic Basis of Disease*, 9th ed. London: Elsevier Health Sciences, 2015.

[2] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, https://doi.org/10.1016/j.media.2017.07.005.

[3] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *Journal of Pathology Informatics*, vol. 7, no. 1, Jan. 2016, Art. no. 29, https://doi.org/10.4103/2153-3539.186902.

[4] D. Komura and S. Ishikawa, "Machine Learning Methods for Histopathological Image Analysis," *Computational and Structural Biotechnology Journal*, vol. 16, pp. 34–42, 2018, https://doi.org/10.1016/j.csbj.2018.01.001.

[5] G. Campanella *et al.*, "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images," *Nature Medicine*, vol. 25, no. 8, pp. 1301–1309, Aug. 2019, https://doi.org/10.1038/s41591-019-0508-1.

[6] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, "Deep Learning for Identifying Metastatic Breast Cancer." arXiv, Jun. 18, 2016, https://doi.org/10.48550/arXiv.1606.05718.

[7] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, Feb. 2020, https://doi.org/10.1007/s11263-019-01228-7.

[8] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, https://doi.org/10.1109/TBME.2015.2496264.

[9] M. Veta *et al.*, "Assessment of algorithms for mitosis detection in breast cancer histopathology images," *Medical Image Analysis*, vol. 20, no. 1, pp. 237–248, Feb. 2015, https://doi.org/10.1016/j.media.2014.11.010.

[10] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer Graphics and Applications*, vol. 21, no. 4, pp. 34–41, Aug. 2001, https://doi.org/10.1109/38.946629.

[11] B. E. Bejnordi *et al.*, "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer," *JAMA*, vol. 318, no. 22, Dec. 2017, Art. no. 2199, https://doi.org/10.1001/jama.2017.14585.

[12] T. Araújo *et al.*, "Classification of breast cancer histology images using Convolutional Neural Networks," *PLoS One*, vol. 12, no. 6, Jun. 2017, Art. no. e0177544, https://doi.org/10.1371/journal.pone.0177544.

[13] H. Wang *et al.*, "Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features," *Journal of Medical Imaging*, vol. 1, no. 3, Oct. 2014, Art. no. 034003, https://doi.org/10.1117/1.JMI.1.3.034003.

[14] M. Ilse, J. M. Tomczak, and M. Welling, "Attention-based Deep Multiple Instance Learning." arXiv, 2018, https://doi.org/10.48550/ARXIV.1802.04712.

[15] V. S. Krushnasamy, O. Al-Omari, A. Sundaram, I. K. Varghese, E. Muniyandy, and M. V. A. L. N. Rao, "LiDAR-Based Climate Change Imaging in Geoscience Using Spatio Extreme Fuzzy Gradient Model," *Remote Sensing in Earth Systems Sciences*, Jan. 2025, https://doi.org/10.1007/s41976-025-00197-5.

[16] L. S. Geetha *et al.*, "Challenges and Solutions in Agile Software Development: A Managerial Perspective on Implementation Practices," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 3, 2025, https://doi.org/10.14569/IJACSA.2025.0160374.

[17] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." arXiv, 2020, https://doi.org/10.48550/ARXIV.2010.11929.

[18] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations." arXiv, 2020, https://doi.org/10.48550/ARXIV.2002.05709.

[19] X. Li, Y. Gu, N. Dvornek, L. H. Staib, P. Ventola, and J. S. Duncan, "Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results," *Medical Image Analysis*, vol. 65, Oct. 2020, Art. no. 101765, https://doi.org/10.1016/j.media.2020.101765.

[20] Y. Ektefaie *et al.*, "Integrative multiomics-histopathology analysis for breast cancer classification," *npj Breast Cancer*, vol. 7, no. 1, Nov. 2021, Art. no. 147, https://doi.org/10.1038/s41523-021-00357-y.

[21] S. M. Fati and M. Alenezi, "Transforming Application Development With Serverless Computing:," *International Journal of Cloud Applications and Computing*, vol. 14, no. 1, pp. 1–16, Dec. 2024, https://doi.org/10.4018/IJCAC.365288.

[22] O. Al-Omari, A. Alyousef, S. Fati, F. Shannaq, and A. Omari, "Governance and Ethical Frameworks for AI Integration in Higher Education: Enhancing Personalized Learning and Legal Compliance," *Journal of Ecohumanism*, vol. 4, no. 2, Jan. 2025, https://doi.org/10.62754/joe.v4i2.5781.

[23] A. Alyousef and O. Al-Omari, "Artificial Intelligence in Healthcare: Bridging Innovation and Regulation," *Journal of Ecohumanism*, vol. 3, no. 8, Jan. 2025, https://doi.org/10.62754/joe.v3i8.5673.

[24] A. Rehman, M. Mujahid, A. Elyassih, B. AlGhofaily, and S. A. O. Bahaj, "Comprehensive Review and Analysis on Facial Emotion Recognition: Performance Insights into Deep and Traditional Learning with Current Updates and Challenges," *Computers, Materials & Continua*, vol. 82, no. 1, pp. 41–72, 2025, https://doi.org/10.32604/cmc.2024.058036.

[25] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Improving sentiment classification using a RoBERTa-based hybrid model," *Frontiers in Human Neuroscience*, vol. 17, Dec. 2023, Art. no. 1292010, https://doi.org/10.3389/fnhum.2023.1292010.

[26] A. Q. A. Hassan, B. B. Al-onazi, M. Maashi, A. A. Darem, I. Abunadi, and A. Mahmud, "Enhancing extractive text summarization using natural language processing with an optimal deep learning model," *AIMS Mathematics*, vol. 9, no. 5, pp. 12588–12609, 2024, https://doi.org/10.3934/math.2024616.

[27] A. Jabbar, S. Iqbal, M. I. Tamimy, A. Rehman, S. A. Bahaj, and T. Saba, "An Analytical Analysis of Text Stemming Methodologies in Information Retrieval and Natural Language Processing Systems," *IEEE Access*, vol. 11, pp. 133681–133702, 2023, https://doi.org/10.1109/ACCESS.2023.3332710.

[28] M. S. Al Huda, T. E. Shrestha, A. Hossain, N. B. Sharif, M. A. Ali, and T. I. Erdei, "DeepMelaNet: Advancing Melanoma Stage Classification in Skin Cancer Diagnosis," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19627–19635, Feb. 2025, https://doi.org/10.48084/etasr.8336.

[29] A. Ammar, A. Koubaa, B. Benjdira, O. Nacar, and S. Sibaee, "Prediction of Arabic Legal Rulings Using Large Language Models," *Electronics*, vol. 13, no. 4, Feb. 2024, Art. no. 764, https://doi.org/10.3390/electronics13040764.

[30] *BreakHis: Breast Cancer Histopathological Image Classification.* (2016), F. A. Spanhol, L. S. Oliveira, C. Petitjean, L. Heutte, C. E. Pokes, F. Negretti and F. S. Rechia. [Online]. Available: http://www.inf.ufpr.br/vri/databases/BreaKHis_v1.tar.gz.

AUTHORS PROFILE

Omaia Al-Omari is an Assistant Professor in the Information Systems Department at Prince Sultan University in Riyadh, Saudi Arabia. He holds a Ph.D. in Computer Science from the National University of Malaysia (UKM) and an M.Sc. and B.Sc. in Computer Information Systems from Yarmouk University. With over 18 years of experience in academia and IT, Dr. Al-Omari has taught at Shaqra University, the University of Ha'il, and Jerash Private University. He has published extensively in prestigious academic journals and has contributed to numerous academic committees. His expertise primarily focuses on artificial intelligence, data science, and information systems.

Omar M.H. Alkhatib is a QC Senior Staff Engineer with extensive expertise in quality control, DevOps transformation, and AI-powered software engineering. He holds a master's degree in computer information systems and has a proven track record of enhancing product quality and operational success through strategic initiatives and innovation. His expertise extends to automation, AI-driven software engineering, and Agile/DevOps methodologies, positioning him as a leader in optimizing development processes and software quality.

Tariq Al-Omari is an Assistant Professor in the Computer Science Department at Jordan University of Science and Technology (JUST) in Jordan. He has over 20 years of industry experience in North America at Intel, IBM, BlackBerry and Morgan Stanley etc. Dr. Al-Omari earned his Ph.D. in Systems and Computer Engineering from Carleton University in 2007. He has publications in prestigious international conferences and journals. His research interests focus on software performance engineering, monitoring, and artificial intelligence.